DEEPESH YADAV

Bangalore, KA
deepeshyadav760@gmail.com
+91- 7275632538

https://github.com/deepeshyadav760 www.linkedin.com/in/deepesh-yadav-b29a29250

EDUCATION

Education Level	Institution	Board/Stream	Year	Score Type	Score
B.Tech	Atria University, Bengaluru	Computer Science (Data Science & ML)	2022 - 2026	CGPA	8.0
Class XII	Vidyagyan School, Bulandshahr	CBSE (PCM)	2021 - 2022	Percentage	94%
Class X	Vidyagyan School, Bulandshahr	CBSE	2019 - 2020	Percentage	96.33%

EXPERIENCE

AI Research Intern – Climate & Sustainability Tech, ACAD 1st May 2025 - present

- Contributing to the development of a Python-based tool for dynamic, location-specific natural resource mapping.
- Assisting in the integration of Google Agricultural Understanding and Google Earth Engine (GEE) for climate action and sustainable agriculture use cases.
- Supporting the optimization of existing APIs and validation of use cases related to AI and decentralized systems

Data Science Intern at Insignia Consultancy Solutions June 2024 - Dec 2024

- During this Internship, I have been working on the NER model.
- I was given tasks for annotating the Resumes and Job descriptions which we will use for our NER model.
- I also have worked on R&D on GenAI and different Machine Learning Algorithms.

Trainee at Intel® Unnati Industrial Training Program - 2024 May 2024 - Jul 2024

- Under this training program I was given a project "Intel® Products Sentiment Analysis from Online Reviews"
- Developed a Python-based automated web scraper to efficiently extract relevant data from Amazon.com.
- Utilized Natural Language Processing (NLP) techniques to analyze sentiment in customer reviews and extract valuable insights.
- Built a machine learning model achieving an accuracy of 85% in predicting sentiment across different review categories.

PROJECTS

Carbon Sequestration Analysis | March 2025

- Conducted geospatial and statistical analysis on carbon sequestration in Sundarbans Mangrove Forest and Chilika Lake using remote sensing data (Landsat 8, SRTM)
- Applied Random Forest and SARIMA models to quantify and forecast CO_2 absorption trends from 2018 to 2024
- Engineered vegetation indices (NDVI, NDWI, MNDVI, EVI) via Google Earth Engine API and Python libraries (scikit-learn, statsmodels, GeoPandas)
- Achieved higher prediction accuracy with SARIMA ($R^2 = 0.66$) by capturing seasonal variations
- Discovered a 1.16% projected decline in Sundarbans carbon stock by 2029 and identified Chilika's high-sequestration zones (24% of lake area)

Time Series Analysis and Forecasting Project from Satellite data using Google Earth Engine (Feb 2025):

- Developed predictive models (SARIMA, Random Forest) to forecast AQI, NDVI, temperature, and precipitation, achieving R^2 of 0.9237 (temperature), 0.8566 (AQI NO_2), 0.6092 (NDVI), and 0.5175 (precipitation) while reducing RMSE.
- Analyzed and processed satellite datasets (Sentinel-2, Sentinel-5P, MODIS, ERA5), identifying key relationships like 0.60 positive correlation between NDVI & precipitation and -0.42 negative correlation between precipitation & AQI (NO_2).

Sentiment Analysis on Ear Wearable Product Reviews Using ML & DL (2023 - 2024)

- Developed a deep learning-based sentiment analysis system for classifying 21,825 Amazon reviews of ear wearable products with 90% accuracy.
- Built an LSTM model with optimized sequential text processing, achieving an F1-score of 0.91 and precision-recall balance of 0.89.
- Automated data scraping using Selenium and implemented a Flask-based UI, providing sentiment predictions with confidence scores for real-time user interaction.

Breast Tumor Classification (2023 - 2024)

- Developed a system analyzing 20,403 breast tumor images across Non-Cancer (8,060), Early Phase (6,133), and Middle Phase (6,210) categories.
- Applied advanced image processing (Gaussian Smoothing, HOG) and tested models:
- Custom CNN (80.6% accuracy - best performer)
- Random Forest (77%), SVM (74%), MobileNetV2 (underperformed).
- Enhances diagnostic efficiency, reducing manual analysis time, supporting early detection, and enabling cost-effective tumor screening.

SKILLS

- **Programming Languages**: Python, SQL, BASH scripting
- **Libraries and Frameworks**: Scikit-Learn, TensorFlow, NLTK, Pandas, NumPy, Seaborn, Matplotlib, Flask, OpenCV, Django, XGBoost LightGBM, GeoPandas, Rasterio, Folium
- **Developer Tools & Databases** :GitHub, mongodb, mysql, MS sql server, RapidMiner AI Studio, GEE, CI/CD pipelines
- **ML Architectures**: Supervised learning, Unsupervised learning, Neural Network , Digital Image Processing, Computer Vision, Time Series Analysis & Forecasting, Remote Sensing, Geospatial Data, Satellite Imaging, LLM, NLP,

CERTIFICATION & ACHIEVEMENTS

- Research paper accepted at IEEE Space
- **1st** position in Data Science Contest by Altair India.
- Machine Learning Professional Certicate by Altair
- Partcipated in "The Beyonders Hackathon" organised by Atria University on 1st and 2nd December 2023
- Machine Learning by IBM, Online
- Python Functions, Files And Dictionaries University Of Michigan, Online
- MySQL by Great Learning, Online
- Attended Art of Computing talks at IISc Benagluru on 17th-18th Feb , 2024